

05-90  
2303

#13



OIKE

## RAW SEQUENCE LISTING

DATE: 02/27/2003

PATENT APPLICATION: US/09/992,095B

TIME: 11:42:09

Input Set : A:\G-091US05DIV-SEQLIST.TXT

Output Set: N:\CRF4\02272003\I992095B.raw

3 <110> APPLICANT: Benjanin, Stephane  
 4 Tanaka, Hiroaki  
 6 <120> TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
 8 <130> FILE REFERENCE: 91.US5.DIV  
 10 <140> CURRENT APPLICATION NUMBER: US 09/992,095B  
 C--> 11 <141> CURRENT FILING DATE: 2003-02-20  
 13 <150> PRIOR APPLICATION NUMBER: US 09/924,340  
 14 <151> PRIOR FILING DATE: 2001-08-06  
 16 <150> PRIOR APPLICATION NUMBER: PCT/IB01/01715  
 17 <151> PRIOR FILING DATE: 2001-08-06  
 19 <150> PRIOR APPLICATION NUMBER: US 60/305,456  
 20 <151> PRIOR FILING DATE: 2001-07-13  
 22 <150> PRIOR APPLICATION NUMBER: US 60/302,277  
 23 <151> PRIOR FILING DATE: 2001-06-29  
 25 <150> PRIOR APPLICATION NUMBER: US 60/298,698  
 26 <151> PRIOR FILING DATE: 2001-06-15  
 28 <150> PRIOR APPLICATION NUMBER: US 60/293,574  
 29 <151> PRIOR FILING DATE: 2001-05-25  
 31 <160> NUMBER OF SEQ ID NOS: 112  
 33 <170> SOFTWARE: JPatent  
 35 <210> SEQ ID NO: 1  
 36 <211> LENGTH: 2016  
 37 <212> TYPE: DNA  
 38 <213> ORGANISM: Homo sapiens  
 40 <220> FEATURE:  
 41 <221> NAME/KEY: 5'UTR  
 42 <222> LOCATION: 1..1434  
 44 <220> FEATURE:  
 45 <221> NAME/KEY: CDS  
 46 <222> LOCATION: 1435..1836  
 48 <220> FEATURE:  
 49 <221> NAME/KEY: 3'UTR  
 50 <222> LOCATION: 1837..2016  
 52 <220> FEATURE:  
 53 <221> NAME/KEY: polyA\_signal  
 54 <222> LOCATION: 1965..1970  
 56 <220> FEATURE:  
 57 <221> NAME/KEY: polyA\_site  
 58 <222> LOCATION: 2001..2016  
 60 <400> SEQUENCE: 1  
 61 aaggtctctc tgcattgcata caccaaggaa aagccacatg aggacataac caggaagaga 60  
 62 gccatcacca agaaccggaa catgcggaca ccctgatctc ggacttctag ccttcagaac 120  
 63 cggttgccaca gttttgatga tcatctctct cccaaccaag atggtggaaa aagcaaaaac 180

ENTERED

## RAW SEQUENCE LISTING

DATE: 02/27/2003

PATENT APPLICATION: US/09/992,095B

TIME: 11:42:09

Input Set : A:\G-091US05DIV-SEQLIST.TXT

Output Set: N:\CRF4\02272003\I992095B.raw

```

64 gtggtgaatc ttggagcaat ccgacaaggc atgaaacgct tccaatttct gttaaactgc 240
65 tgtgagccag ggacaattcc tgatgcctcc atcctagcag ctgccttgga tctactatgc 300
66 ggcattcttc tgattcattt ttctccattt gtgctgtttt tctctgtgat gtgaatccat 360
67 ccctatccat tatgtcatgc ctccatcttt tgctgttctt tcagattgca ctgagccata 420
68 agaggaagcc cctgtggtgg ccagagcagc cttgttcctg gaatgtgctc gttttgttca 480
69 ccgctgcaac cgtggcaact ggccagagtg gatgaaaggg caccacgtga acatcaccaa 540
70 gaaaggactt tcccggggac gctctcccat tgtgggcaac aagcgaaacc agaagctgca 600
71 gtggaatgca gccaaagctct tctaccaatg gggagacaag gaaaaaagggt gaagaataaa 660
72 aggaaattca agaggaccaa gtttctgcta attttagaca gagctgaaca taaacacaca 720
73 taaagagggt ccatatatct ctcttttctt aaagattact tggaataact gttacaattt 780
74 ccgttaataa ttcaagctgaa tgtgtctacc aatgtgctta ccaactaagg caattggcgt 840
75 ccgattgaat gagctgtgcc acggggaaaag tgagagccca gccaacctgc tgggtctcat 900
76 ttacgatgaa gagaccaaga ggagacttag aaaggaggat gaggaggaag actttttaga 960
77 tgacattcca ctttcaagtc aatacacagc tcatcttgca tttaaaagct gattatggtg 1020
78 caagcaactt tcgggctgga aattctacag aagcttgtct tttccattct tgatgagagg 1080
79 caaagtcccc ggcaacaaat taactcagga gagaaaatgg ttttctgaa aaaaacgata 1140
80 gcttaaatat ctacagaaag accgtaattt ccacctattt tcaaatgaaa tcgtgaaaaa 1200
81 cacatttgga ctagagctga aacaacttca ctgccctcaa aacagcaaga cagacatccc 1260
82 tcataaaatg aactgacaga atttttatag ctccaaatct agttcactgc catatacata 1320
83 gtctaaatct gattgaatag cagcgtagaa atcttgcgaa attacttccc atttctgttt 1380
84 tcgttaaaaag gtactgtgaa cccctctaaa tgcggttgcc cctttgcctt gaag atg 1437
85 Met
86 1
87 gca gca tgt cag ctt ctt ctg gag att acc acc ttc ctg cga gag acc 1485
88 Ala Ala Cys Gln Leu Leu Leu Glu Ile Thr Thr Phe Leu Arg Glu Thr
89 5 10 15
90 ttt tct tgc ctg ccc aga cct cgc act gag cct ctg gtg gct tca acg 1533
91 Phe Ser Cys Leu Pro Arg Pro Arg Thr Glu Pro Leu Val Ala Ser Thr
92 20 25 30
93 gac cac acc aaa atg cca tct caa atg gaa cac gcc atg gaa acc atg 1581
94 Asp His Thr Lys Met Pro Ser Gln Met Glu His Ala Met Glu Thr Met
95 35 40 45
96 atg ttt aca ttt cac aaa ttc gct ggg gat aaa ggc tac tta aca aag 1629
97 Met Phe Thr Phe His Lys Phe Ala Gly Asp Lys Gly Tyr Leu Thr Lys
98 50 55 60 65
99 gag gac ctg aga gta ctc atg gaa aag gag ttc cct gga ttt ttg gaa 1677
100 Glu Asp Leu Arg Val Leu Met Glu Lys Glu Phe Pro Gly Phe Leu Glu
101 70 75 80
102 aat caa aaa gac cct ctg gct gtg gac aaa ata atg aag gac ctg gac 1725
103 Asn Gln Lys Asp Pro Leu Ala Val Asp Lys Ile Met Lys Asp Leu Asp
104 85 90 95
105 cag tgt aga gat ggc aaa gtg ggc ttc cag agc ttc ttt tcc cta att 1773
106 Gln Cys Arg Asp Gly Lys Val Gly Phe Gln Ser Phe Phe Ser Leu Ile
107 100 105 110
108 gcg ggc ctc acc att gca tgc aat gac tat ttt gta gta cac atg aag 1821
109 Ala Gly Leu Thr Ile Ala Cys Asn Asp Tyr Phe Val Val His Met Lys
110 115 120 125
111 cag aag gga aag aag taggcagaaa tgagcagttc gctcctccct gataagagtt 1876
112 Gln Lys Gly Lys Lys

```

## RAW SEQUENCE LISTING

DATE: 02/27/2003

PATENT APPLICATION: US/09/992,095B

TIME: 11:42:09

Input Set : A:\G-091US05DIV-SEQLIST.TXT

Output Set: N:\CRF4\02272003\I992095B.raw

```

113 130
114 gtcccaaagg gtcgcttaag gaatctgccc cacagcttcc cccatagaag gatttcatga 1936
115 gcagatcagg acacttagca aatgtaaaaa taaaatctaa ctctcatttg acaagcagag 1996
116 aaagaaaaaa aaaaaaaaat                                     2016
118 <210> SEQ ID NO: 2
119 <211> LENGTH: 134
120 <212> TYPE: PRT
121 <213> ORGANISM: Homo sapiens
123 <400> SEQUENCE: 2
124 Met Ala Ala Cys Gln Leu Leu Leu Glu Ile Thr Thr Phe Leu Arg Glu
125 1 5 10 15
126 Thr Phe Ser Cys Leu Pro Arg Pro Arg Thr Glu Pro Leu Val Ala Ser
127 20 25 30
128 Thr Asp His Thr Lys Met Pro Ser Gln Met Glu His Ala Met Glu Thr
129 35 40 45
130 Met Met Phe Thr Phe His Lys Phe Ala Gly Asp Lys Gly Tyr Leu Thr
131 50 55 60
132 Lys Glu Asp Leu Arg Val Leu Met Glu Lys Glu Phe Pro Gly Phe Leu
133 65 70 75 80
134 Glu Asn Gln Lys Asp Pro Leu Ala Val Asp Lys Ile Met Lys Asp Leu
135 85 90 95
136 Asp Gln Cys Arg Asp Gly Lys Val Gly Phe Gln Ser Phe Phe Ser Leu
137 100 105 110
138 Ile Ala Gly Leu Thr Ile Ala Cys Asn Asp Tyr Phe Val Val His Met
139 115 120 125
140 Lys Gln Lys Gly Lys Lys
141 130
143 <210> SEQ ID NO: 3
144 <211> LENGTH: 1081
145 <212> TYPE: DNA
146 <213> ORGANISM: Homo sapiens
148 <220> FEATURE:
149 <221> NAME/KEY: 5'UTR
150 <222> LOCATION: 1..38
152 <220> FEATURE:
153 <221> NAME/KEY: CDS
154 <222> LOCATION: 39..917
156 <220> FEATURE:
157 <221> NAME/KEY: 3'UTR
158 <222> LOCATION: 918..1081
160 <220> FEATURE:
161 <221> NAME/KEY: polyA_signal
162 <222> LOCATION: 1045..1050
164 <220> FEATURE:
165 <221> NAME/KEY: polyA_site
166 <222> LOCATION: 1066..1081
168 <400> SEQUENCE: 3
169 gtccagcctg ttgctgatgc tgccgtgcgg tacttgtc atg gag ctg gca ctg cgg 56
170 Met Glu Leu Ala Leu Arg

```



## RAW SEQUENCE LISTING

DATE: 02/27/2003

PATENT APPLICATION: US/09/992,095B

TIME: 11:42:09

Input Set : A:\G-091US05DIV-SEQLIST.TXT

Output Set: N:\CRF4\02272003\I992095B.raw

```

220 gta aaa agg gga aac act cag agg cta gcc tgc ttg gct ttt tct ggt      872
221 Val Lys Arg Gly Asn Thr Gln Arg Leu Ala Cys Leu Ala Phe Ser Gly
222      240      245      250
223 ggg tac agg gcc cat ggt tgg tgt tgt caa act tgg agt cta cac      917
224 Gly Tyr Arg Ala His Gly Trp Cys Cys Gln Thr Trp Ser Leu His
225      255      260      265
226 tgaggctccc cacatatctg caaatgattg catgctggat aataaatctc ttgggtctaa  977
227 gcagtgatgt agtggctcct tacagagtca gaaagccacc caggcctgca agacttgctt 1037
228 gtccttcact aaatgtatgg attctattaa aaaaaaaaaa aaaa      1081
230 <210> SEQ ID NO: 4
231 <211> LENGTH: 293
232 <212> TYPE: PRT
233 <213> ORGANISM: Homo sapiens
235 <220> FEATURE:
236 <221> NAME/KEY: SIGNAL
237 <222> LOCATION: 1..26
239 <400> SEQUENCE: 4
240 Met Glu Leu Ala Leu Arg Arg Ser Pro Val Pro Arg Trp Leu Leu Leu
241      -25      -20      -15
242 Leu Pro Leu Leu Leu Gly Leu Asn Ala Gly Ala Val Ile Asp Trp Pro
243      -10      -5      1      5
244 Thr Glu Glu Gly Lys Glu Val Trp Asp Tyr Val Thr Val Arg Lys Asp
245      10      15      20
246 Ala Tyr Met Phe Trp Trp Leu Tyr Tyr Ala Thr Asn Ser Cys Lys Asn
247      25      30      35
248 Phe Ser Glu Leu Pro Leu Val Met Trp Leu Gln Gly Gly Pro Gly Gly
249      40      45      50
250 Ser Ser Thr Gly Phe Gly Asn Phe Glu Glu Ile Gly Pro Leu Asp Ser
251      55      60      65      70
252 Asp Leu Lys Pro Arg Lys Thr Thr Trp Leu Gln Ala Ala Ser Leu Leu
253      75      80      85
254 Phe Val Asp Asn Pro Val Gly Thr Gly Phe Ser Tyr Val Asn Gly Ser
255      90      95      100
256 Gly Ala Tyr Ala Lys Asp Leu Ala Met Val Ala Ser Asp Met Met Val
257      105      110      115
258 Leu Leu Lys Thr Phe Phe Ser Cys His Lys Glu Phe Gln Thr Val Pro
259      120      125      130
260 Phe Tyr Ile Phe Ser Glu Ser Tyr Gly Gly Lys Met Ala Ala Gly Ile
261      135      140      145      150
262 Gly Leu Glu Leu Tyr Lys Ala Ile Gln Arg Gly Thr Ile Lys Cys Asn
263      155      160      165
264 Phe Ala Gly Val Ala Leu Gly Asp Ser Trp Ile Ser Pro Val Asp Ser
265      170      175      180
266 Val Leu Ser Trp Gly Pro Tyr Leu Tyr Ser Met Ser Leu Leu Glu Asp
267      185      190      195
268 Lys Gly Leu Ala Glu Val Ser Lys Val Ala Glu Gln Val Leu Asn Ala
269      200      205      210
270 Val Asn Lys Gly Leu Tyr Arg Glu Ala Thr Glu Leu Trp Gly Lys Ala
271      215      220      225      230

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/09/992,095B

DATE: 02/27/2003  
TIME: 11:42:10

Input Set : A:\G-091US05DIV-SEQLIST.TXT  
Output Set: N:\CRF4\02272003\I992095B.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:27; Xaa Pos. 116,233  
Seq#:28; Xaa Pos. 116,233  
Seq#:29; Xaa Pos. 30  
Seq#:30; Xaa Pos. 29  
Seq#:71; Xaa Pos. 157  
Seq#:72; Xaa Pos. 156

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/09/992,095B

DATE: 02/27/2003

TIME: 11:42:10

Input Set : A:\G-091US05DIV-SEQLIST.TXT

Output Set: N:\CRF4\02272003\I992095B.raw

L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:1487 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:27 after pos.:502  
M:341 Repeated in SeqNo=27  
L:1554 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:28 after pos.:112  
M:341 Repeated in SeqNo=28  
L:1614 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:29 after pos.:500  
L:1677 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30 after pos.:48  
L:4116 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:71 after pos.:529  
L:4183 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:72 after pos.:160